5

5.1

5.2

5.3

5.4

5.5

5.1

5.1

- ❑

定义5.1 $D \subset \mathbf{R}^n$, $x, y \in D$

$$\lambda x + (1-\lambda) y \in D \quad \forall\, 0 \le \lambda \le 1$$

D

- ❑ :
  - ➢ $H = \{x | p^{\mathrm{T}} x = \alpha\}$ $p \in \mathbf{R}^n$
    α
  - ➢ $H^- = \{x | p^{\mathrm{T}} x \le \alpha\}$ $H^+ = \{x | p^{\mathrm{T}} x \ge \alpha\}$
    $H_0^- = \{x | p^{\mathrm{T}} x < \alpha\}$ $H_0^+ = \{x | p^{\mathrm{T}} x > \alpha\}$
    - ❖ $D = \{x | p_i^{\mathrm{T}} x \le \beta_i, i = 1, \cdots, m\}$
      (polyhetral) p_i β_i

5.1

❑

$D_1, D_2 \subset \mathbf{R}^n$

(1) $D_1 \cap D_2 = \{x | x \in D_1 \quad x \in D_2\}$

(2) $D_1 + D_2 = \{x + y | x \in D_1, y \in D_2\}$

(3) $D_1 - D_2 = \{x - y | x \in D_1, y \in D_2\}$

(4) α $\alpha D_1 = \{\alpha x | x \in D_1\}$

定理5.1 $D \subset \mathbf{R}^n$ D m $x^{(i)}(i = 1,2,\cdots,m)$ D

$$\sum_{i=1}^{m} \alpha_i x^{(i)} \in D, \alpha_i \geq 0(i = 1,2,\cdots,m) \quad \sum_{i=1}^{m} \alpha_i = 1$$

5.1

定义5.2 $D_1, D_2 \subset \mathbf{R}^n$ $\alpha \in \mathbf{R}^n$ β

$$D_1 \subset H^+ = \{x \in \mathbf{R}^n | \alpha^{\mathrm{T}} x \geq \beta\}$$

$$D_2 \subset H^- = \{x \in \mathbf{R}^n | \alpha^{\mathrm{T}} x \leq \beta\}$$

$$H = \{x \in \mathbf{R}^n | \alpha^{\mathrm{T}} x = \beta\}$$

D_1 D_2

$$D_1 \subset H_0^+ = \{x \in \mathbf{R}^n | \alpha^{\mathrm{T}} x > \beta\}$$

$$D_2 \subset H_0^- = \{x \in \mathbf{R}^n | \alpha^{\mathrm{T}} x < \beta\}$$

H D_1 D_2 H_0^+ H_0^- H^+ H^-

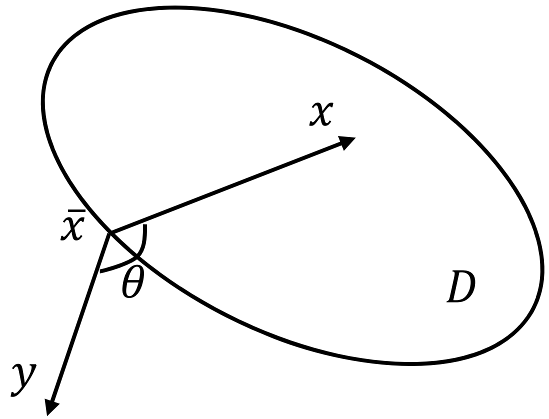
5.1

定理5.2 ( ) $D \subset \mathbf{R}^n$ $y \in \mathbf{R}^n$ $y \notin D$

(1) $\bar{x} \in D$ D y

$$\|\bar{x} - y\| = \inf_{x \in D} \|x - y\|$$

(2)$\bar{x} \in D$ y D



$$(x - \bar{x})^{\mathrm{T}}(\bar{x} - y) \geq 0, \forall x \in D$$

$$\langle x - \bar{x}, y - \bar{x} \rangle \leq 0, \forall x \in D$$

定理5.3 $D \subset \mathbf{R}^n$ $y \in \mathbf{R}^n$ $y \notin D$

$\alpha \in \mathbf{R}^n$ β

$$\alpha^{\mathrm{T}} x \leq \beta < \alpha^{\mathrm{T}} y \quad \forall x \in D$$

$$H = \{x \in \mathbf{R}^n | \alpha^{\mathrm{T}} x = \beta\}$$ y D

5.1

❑

定义5.3 $f(x)$ D x $y \in D$ $\lambda \in [0,1]$

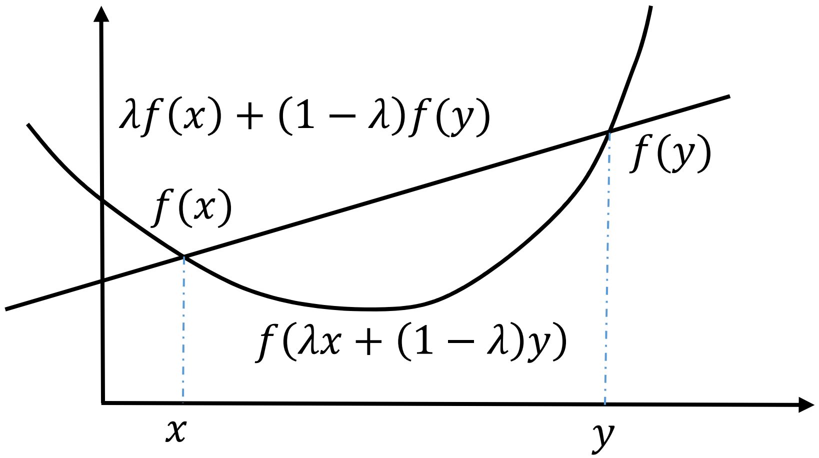
$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$$

$f(x)$ D

x $y \in D, x \neq y$ $\lambda \in (0,1)$

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y)$$

$f(x)$ D



5.1

❑

(1) f D $\alpha \geq 0$ αf

D

(2) f_1 f_2 D $f_1 + f_2$

D

(3) $f_i(x)(i = 1, \cdots, m)$ D

$f(x) = \max\limits_{1 \leq i \leq m} |f_i(x)|$ D

(4) $f_i(x)(i = 1, \cdots, m)$ D

$f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$ D $\alpha_i \geq 0(i = 1, \cdots, m)$

5.1

定理5.4 x^*

(1) x^*

(2) x^*

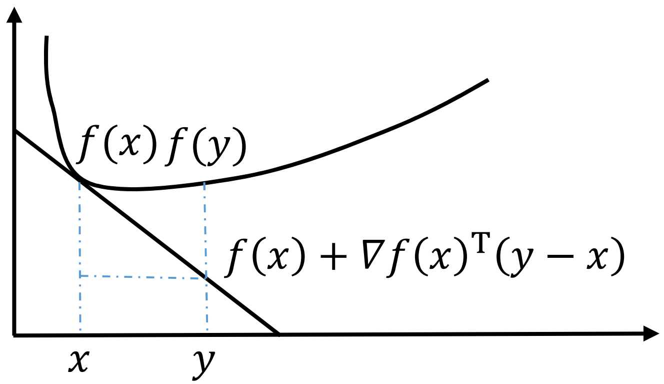
定理5.5 $f(x)$ D

(1) $f(x)$ D

$$f(y) \geq f(x) + \nabla f(x)^{\mathrm{T}}(y-x), \forall x, y \in D$$

(2) $f(x)$ D

$$f(y) > f(x) + \nabla f(x)^{\mathrm{T}}(y-x), \forall x, y \in D, x \neq y$$



5.1

定理5.6 $f(x)$ $D \subset \mathbf{R}^n$

(1) $f(x)$ D $f(x)$ Hesse (

) $\nabla^2 f(x)$ D $x \in D$,

$$y^{\mathrm{T}} \nabla^2 f(x) y \geq 0, \forall \mathrm{y} \in \mathbf{R}^n$$

(2) $f(x)$ Hesse $\nabla^2 f(x)$ D $f(x)$ D

$f(x)$ D $\nabla^2 f(x)$ D

5.2

❑

$$\mathbf{P:}\ \min\ f(x) \qquad (5.1)\ (\qquad)$$

$$s.t.\ h_i(x) = 0, i = 1,2,\cdots,m \qquad (5.2)\ (\qquad)$$

$$h_i(x) \geq 0, i = m+1,\cdots,p \qquad (5.3)\ (\qquad)$$

x n

❑ ,

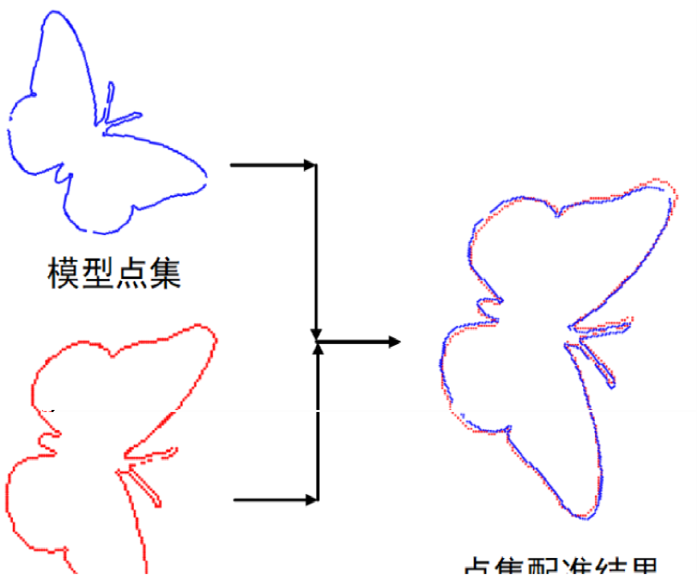
$\min f(x)$

定义5.4 ( ) (5.2) (5.3) x ,

(Feasible Point)

5.2

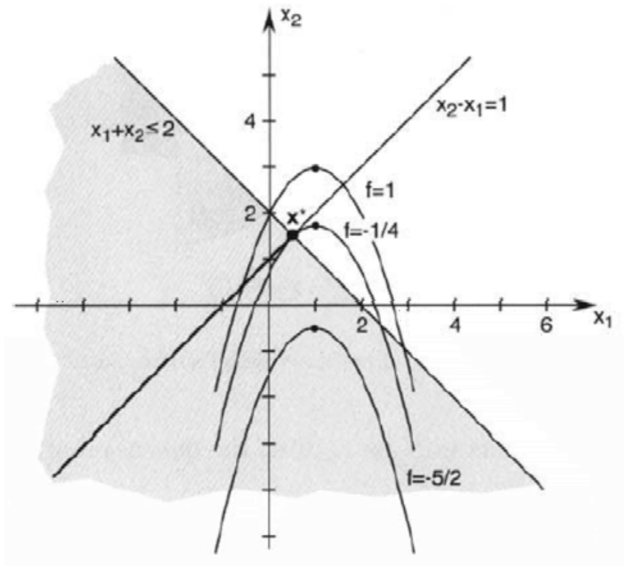
-



5.2

例5.1

$$\text{minimize} \quad (x_1-1)^2+x_2-2$$
$$\text{subject to} \quad x_2-x_1=1$$
$$x_1+x_2\le 2$$



- $f(x_1,x_2)=(x_1-1)^2+x_2-2$

  $h(x_1,x_2)=x_2-x_1-1$ $g(x_1,x_2)=x_1+x_2-2$

- f

- $f=-1/4$

  $x^*=[\frac{1}{2},\frac{3}{2}]^T$

5.2

定义5.5（ ） Feasible Region) D :

$$D = \{x | h_i(x) = 0, i = 1,\cdots,m, h_j(x) \geq 0, i = m+1,\cdots,p, x \in \mathbf{R}^n\}$$

- $h_i(x)$ D

定义5.6（ ） $x^* \in D$ $x \in D$ $f(x^*) \leq f(x)$ x^* (P)

- $x^* \in D$ $x \neq x^*$ $f(x^*) < f(x)$ x^* (P)

5.2

定义5.7 ( ) $x^* \in D$ x^* $N_\varepsilon(x^*)$

$x \in D \cap N_\varepsilon(x^*)$ $f(x^*) \le f(x)$

(P)

- $N_\varepsilon(x^*) = \{x \mid \|x - x^*\| < \varepsilon, \varepsilon > 0\}$
- $x \neq x^*$ x^* (P)

-

- (P) $f(x)$ (5.2)

  (5.3) D

-

5.2

5.2

❑ ( )

定义5.8 n $f: \mathbf{R}^n \to \mathbf{R}$ $\Omega \subset \mathbf{R}^n$.

Ω x^* $\varepsilon > 0$

$\|x - x^*\| < \varepsilon, x \in \Omega \backslash \{x^*\}$ x $f(x) \geq f(x^*)$

x^* f Ω

$x \in \Omega \backslash \{x^*\}$ $f(x) \geq f(x^*)$ x^* f

Ω

➢ $f(x) \geq f(x^*)$ $f(x) > f(x^*)$

5.2

定义5.9 $f(x)$ $\mathbf{R}^n$ $\bar{x} \in \mathbf{R}^n$

$s \in \mathbf{R}^n$ $\delta > 0$

$$f(\bar{x} + \alpha s) < f(\bar{x}), \forall \alpha \in (0, \delta)$$

s $f(x)$ $\bar{x}$ $\bar{x}$

$D(\bar{x})$.

定理5.7 $f(x)$ $\bar{x}$ $s \in \mathbf{R}^n$

$$\nabla f(\bar{x})^{\mathrm{T}} s < 0$$

s $f(x)$ $\bar{x}$.

定理5.8 ( ) $f\colon \mathrm{D} \subset \mathbf{R}^n \rightarrow \mathbf{R}^1$ D

$x^* \in \mathrm{D}$ $\min\limits_{x \in \mathbf{R}^n} f(x)$

$$g(x^*) = \nabla f(x^*) = 0$$

5.2

定理5.9 ( ) $f: \mathrm{D} \subset \mathbf{R}^n \to \mathbf{R}^1$ D

$x^* \in \mathrm{D}$ $\min_{x \in \mathbf{R}^n} f(x)$

$$g(x^*) = 0, G(x^*) = \nabla^2 f(x^*) \geq 0$$

定理5.10 ( ) $f: \mathrm{D} \subset \mathbf{R}^n \to \mathbf{R}^1$ D

$x^* \in D$ f

$$g(x^*) = 0 \quad G(x^*)$$

定理5.11 ( ) $f: \mathrm{D} \subset \mathbf{R}^n \to \mathbf{R}^1$ $f \in C^1$.

x^* $g(x^*) = 0$

5.3

❑

(1) $x^{(0)}$ $k = 0$;

(2) $x^{(k)}$;

(3) $x^{(k)}$ $s^{(k)}$;

(4) $x^{(k+1)} = x^{(k)} + s^{(k)}$

$k = k + 1$ (2).

5.3

- ❑
  - ➢ $x^{(0)}$ x^* x^*
  - ➢ $x^{(0)} \in D$ x^*
- ❑ x^*
- ❑

5.3

❑

定义5.9 $\{x^{(k)}\} \subset \mathbf{R}^n$ x^*

$$e_k = x^{(k)} - x^*$$

C r

$$\lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|^r} = C$$

$\{x^{(k)}\}$ r x^*(C)

❑ $r = 1, 0 < C < 1$

$$\|e_{k+1}\| \leq C\|e_k\|$$

❑ $r = 1, C = 0$ $\{x^{(k)}\}$ x^*

$r > 1$

5.3

❑

➢

(1) $\left\|x^{(k+1)} - x^{(k)}\right\| \leq \varepsilon$ $\quad \dfrac{\left\|x^{(k+1)} - x^{(k)}\right\|}{\left\|x^{(k)}\right\|} \leq \varepsilon$

(2) $\left|f\left(x^{(k+1)}\right) - f(x^{(k)})\right| \leq \varepsilon$ $\quad \dfrac{\left|f\left(x^{(k+1)}\right) - f(x^{(k)})\right|}{\left|f(x^{(k)})\right|} \leq \varepsilon$

(3) $\left\|\nabla f(x^{(k)})\right\| \leq \varepsilon$

(4) .

$\varepsilon > 0$

5.3

- ❑

$$\begin{cases} \begin{cases} Fibonacci \end{cases} \\ \begin{cases} Goldstein \\ Wolfe \\ Armijo \end{cases} \end{cases}$$

Â i • c

‰ð P ã c

Â Î Ã Ý

¾— m â 4 › µ ÙÛ ì ÷ \ « • > = ák>—?

î ( )Ù (LTÞ IB ÙÞ)@

¨ > = á ©?Ý á ƒ

¾¨ \ « • # H + f µ ãk ä ã O ä

¾* A û µ Ô á ƒ › î (ã) y î (ä) Ý 4 Û Ð í ”

X ! • > = ákã © X ! • : ä á > ?

¾í ” e • • x † © = • x † •ož • • m

= • + f µ %² µ

¾ ý Q ¬ } Ç # H%² + f µ k w ï • + f µ Ô

á ƒ › Ð Ò A ûož M Ö i k Â e G ß â 4 µ ÙÛ

5.3

❑

(1) $[a_0, b_0]$ $\delta > 0$

λ_0, μ_0

$$\lambda_0 = a_0 + 0.382(b_0 - a_0)$$

$$\mu_0 = a_0 + 0.618(b_0 - a_0)$$

$\varphi(\lambda_0)$ $\varphi(\mu_0)$ $k = 0$

(2) $\varphi(\lambda_k) > \varphi(\mu_k)$ (3) (4)

(3) $b_k - \lambda_k \le \delta$ μ_k

$$a_{k+1} \coloneqq \lambda_k, b_{k+1} \coloneqq b_k, \lambda_{k+1} \coloneqq \mu_k$$

$$\varphi(\lambda_{k+1}) \coloneqq \varphi(\mu_k), \mu_{k+1} \coloneqq a_{k+1} + 0.618(b_{k+1} - a_{k+1})$$

$\varphi(\lambda_{k+1})$ (5)

5.3

(4) $\mu_k - a_k \le \delta$ λ_k

$$a_{k+1} \coloneqq a_k, b_{k+1} \coloneqq \mu_k, \mu_{k+1} \coloneqq \lambda_k$$

$$\varphi(\mu_{k+1}) \coloneqq \varphi(\lambda_k)\,, \lambda_{k+1} \coloneqq a_{k+1} + 0.382(b_{k+1} - a_{k+1})$$

$\varphi(\lambda_{k+1})$ (5)

(5) $k \coloneqq k + 1$ (2)

5.3

-

  - $[a_1, b_1]$ k $[a_k, b_k]$

    $$\varphi'(a_k) \leq 0, \varphi'(b_k) \geq 0$$

  - $c_k = \frac{1}{2}(a_k + b_k)$ $\varphi'(c_k) \geq 0$ $a_{k+1} = a_k$, $b_{k+1} = c_k$
  - $\varphi'(c_k) \leq 0$ $a_{k+1} = c_k$, $b_{k+1} = b_k$

    $[a_{k+1}, b_{k+1}]$
  -
  -

    $\frac{1}{2}$.
  - δ n

    $$(\frac{1}{2})^n \leq \frac{\delta}{b_1 - a_1}$$

5.3

❑

(1) $[a_1, b_1]$ δ $k = 1$

(2) $c_k = \frac{1}{2}(a_k + b_k)$ $\varphi'(c_k)$

$\varphi'(c_k) = 0$ c_k

$\varphi'(c_k) > 0$ (3)

$\varphi'(c_k) < 0$ (4)

(3) $a_{k+1} = a_k$, $b_{k+1} = c_k$ (5)

(4) $a_{k+1} = c_k$, $b_{k+1} = b_k$ (5)

(5) $b_{k+1} - a_{k+1} \leq \delta$

$k := k + 1$ (1)

5.4

-

$$\min f(x) \qquad\qquad f: \mathbf{R}^n \to \mathbf{R}^1$$

-

  -

$$f(x) = 0$$

$$\Updownarrow$$

$$\min_x \|f(x) - 0\|^2$$

  -

    Hesse

5.4

❑

$Fibonacci$

5.4

-

- $f(x)$ x_k $g_k \triangleq \nabla f(x_k) \neq 0$ $f(x)$ x_k Taylor ,

$$f(x) = f(x_k) + g_k^{\mathrm{T}}(x - x_k) + o(\|x - x_k\|)$$

- $x - x_k = \alpha d_k$

$$f(x_k + \alpha d_k) = f(x_k) + \alpha g_k^{\mathrm{T}} d_k + o(\|\alpha d_k\|)$$

- d_k $g_k^{\mathrm{T}} d_k < 0$ d_k $f(x_k + \alpha d_k) <$ $f(x_k)$ f d d

$$\min_{d}\ g_k^{\mathrm{T}} d$$
$$s.t.\ \|d\| = 1$$

5.4

- $g_k^{\mathrm{T}} d = -\|g_k\| \|d\| cos\theta_k = -\|g_k\| cos\theta_k, \quad \theta_k = 0, d = \frac{g_k}{\|g_k\|}$
-

$$x_{k+1} = x_k - \alpha_k g_k$$

α_k

-

(1) $x_0 \in \mathbf{R}^n$ $0 \leq \varepsilon << 1$

(2) $d_k = -g_k$ $\|-g_k\| \leq \varepsilon$

(3) α_k

(4) $x_{k+1} = x_k + \alpha_k d_k$

(5) $k := k + 1$ 2

5.4

❑

$f(x)$ x_k Taylor

- $f(x)$ $x_k \in \mathbf{R}^n$ Hesse $\nabla^2 f(x_k)$ x_k Taylor f

$$f(x_k + s) \approx q^{(k)}(s) = f(x_k) + \nabla f(x_k)^{\mathrm{T}} s + \frac{1}{2} s^{\mathrm{T}} \nabla^2 f(x_k) s$$

$s = x - x_k$, $q^{(k)}(s)$ $f(x)$

$$\nabla q^{(k)}(s) = \nabla f(x_k) + \nabla^2 f(x_k) s = 0$$

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$$

5.4

- ❑ $G_k \triangleq \nabla^2 f(x_k)$ $g_k \triangleq \nabla f(x_k)$ $x_{k+1} = x_k - G_k^{-1} g_k$
- ❑

(1) $x_0,$ $\varepsilon > 0$ $k := 0$

(2) g_k $\|g_k\| \leq \varepsilon$ x_k

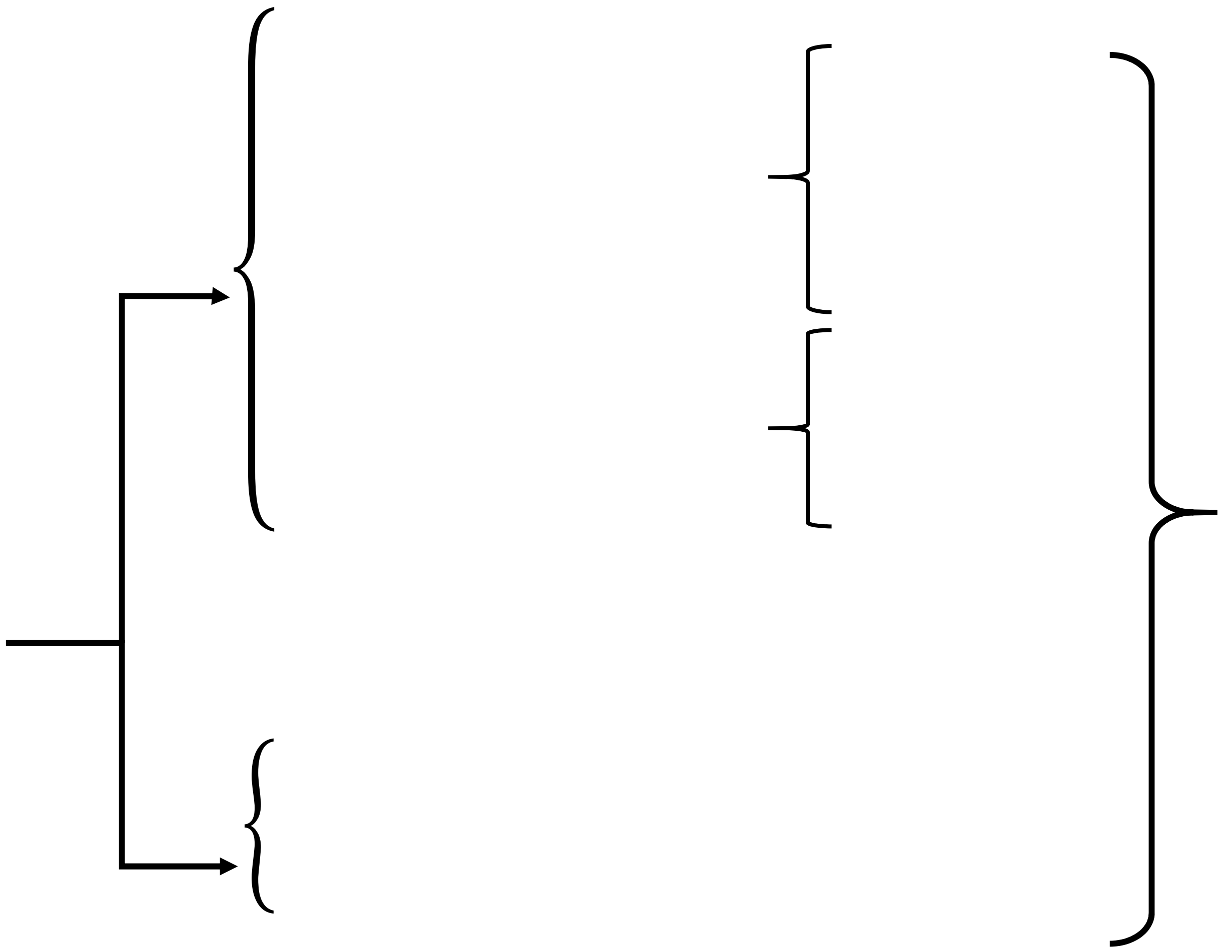
(3) $G_k d = -g_k$ d_k

(4) α_k

$$f(x_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(x_k + \alpha d_k)$$

(5) $x_{k+1} = x_k + \alpha_k d_k, k := k + 1$ (2)

5.5



5.5

” ”

5.5

-

  -

$$\min \quad q(x) = \frac{1}{2} x^{\mathrm{T}} G x + g^{\mathrm{T}} x \qquad (5.4)$$

$$s.t. \quad a_i^{\mathrm{T}} k = b_i \qquad i \in \mathcal{E} \qquad (5.5)$$

$$a_i^{\mathrm{T}} k \geq b_i \qquad i \in \mathcal{J} \qquad (5.6)$$

G $n \times n$ ε τ

g x $i \in \mathcal{E} \cup \mathcal{J}$ n

5.5

-

$$\begin{aligned} &\text{minimize} && f(\boldsymbol{x}) \\ &\text{subject to} && \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0} \end{aligned}$$

$\boldsymbol{x} \in \mathbb{R}^n, f: \mathbb{R}^n \to \mathbb{R}, \boldsymbol{h}: \mathbb{R}^n \to \mathbb{R}^m, \boldsymbol{h} = [h_1, \dots, h_m]^T, m \leq n$

- $\boldsymbol{h}$ $\boldsymbol{h} \in \mathcal{C}^1$

定义5.11 $h_1(\boldsymbol{x}^*) = 0, \dots, h_m(\boldsymbol{x}^*) = 0$ $\boldsymbol{x}^*$

$\nabla h_1(\boldsymbol{x}^*), \dots, \nabla h_m(\boldsymbol{x}^*)$ $\boldsymbol{x}^*$

5.5

- $D\boldsymbol{h}(\boldsymbol{x}^*)$ $\boldsymbol{h}=[h_1,\dots,h_m]^T$ $\boldsymbol{x}^*$

$$D\boldsymbol{h}(\boldsymbol{x}^*)=\begin{bmatrix} Dh_1(\boldsymbol{x}^*) \\ \vdots \\ Dh_m(\boldsymbol{x}^*) \end{bmatrix}=\begin{bmatrix} \nabla h_1(\boldsymbol{x}^*)^{\boldsymbol{T}} \\ \vdots \\ \nabla h_m(\boldsymbol{x}^*)^{\boldsymbol{T}} \end{bmatrix}$$

$$\mathrm{rank} D\boldsymbol{h}(\boldsymbol{x}^*)=m$$

$\boldsymbol{x}^*$是

- $h_1(\boldsymbol{x})=0,\dots,h_m(\boldsymbol{x})=0, h_i\colon\mathbb{R}^n\to\mathbb{R}$

$$S=\{\boldsymbol{x}\in\mathbb{R}^n\colon h_1(\boldsymbol{x})=0,\dots,h_m(\boldsymbol{x})=0\}$$

- S S $n-m$

5.5

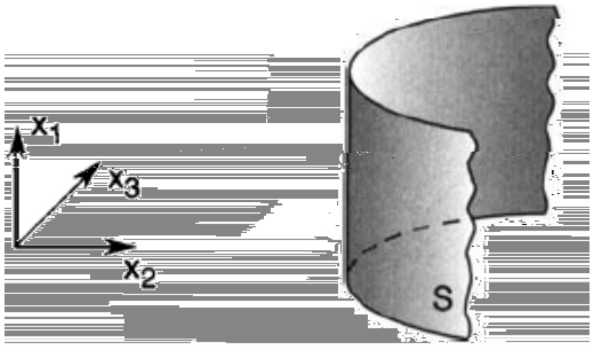
例5.2 $n=3$ $m=1$ $\mathbb{R}^3$ S

S

$$h_1(\boldsymbol{x}) = \boldsymbol{x_2} - \boldsymbol{x_3^2} = \boldsymbol{0}$$

$\nabla h_1(\boldsymbol{x}) = [0,1,-2x_3]^{\mathrm{T}}$ $x \in \mathbb{R}^3$ $\nabla h_1(\boldsymbol{x}) \neq \boldsymbol{0}$

$$dimS = \dim\{\boldsymbol{x}: h_1(\boldsymbol{x}) = 0\} = n - m = 2$$



$S=\{[x_1,x_2,x_3]^{\mathrm{T}} : x_2 - x_3^2 = 0\}$

5.5

定理5.12 （拉格朗日定理） $\boldsymbol{x}^*$ $f:\mathbb{R}^n\to\mathbb{R}$

$\boldsymbol{h}(\boldsymbol{x})=0$ $\boldsymbol{h}:\mathbb{R}^n\to\mathbb{R}^m,\ m\le n,$ $\boldsymbol{x}^*$

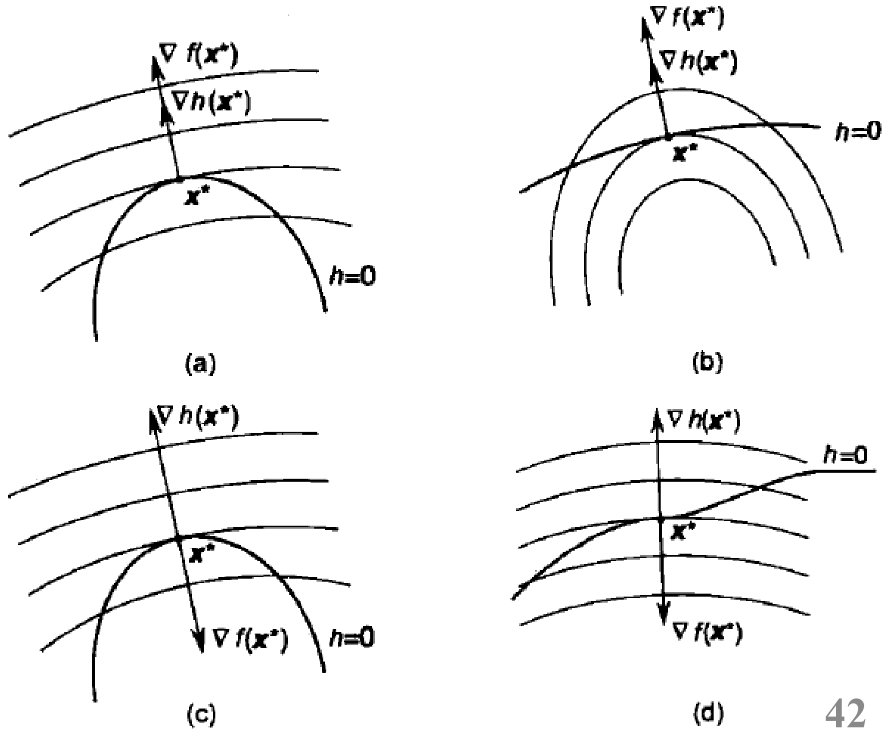
$\lambda^*\in R^m$

$$Df(\boldsymbol{x}^*)+\lambda^{*T}D\boldsymbol{h}(\boldsymbol{x}^*)=0^T$$

❑

$\nabla f(x^*)+\lambda^*\nabla h(x^*)$ $=0$

$\nabla h(x^*)=0$



5.5

例5.3

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & h(x) = 0 \end{aligned}$$

$f(x) = x,$

$$h(x) = \begin{cases} x^2 & x < 0 \\ 0 & 0 \le x \le 1 \\ (x-1)^2 & x > 1 \end{cases}$$

$[0,1]$ $x^* = 0$ $f'(x^*) =$ $1, h'(x^*) = 0$ x^*

- ❑ x^*

5.5

- $f: \mathbb{R}^n \to \mathbb{R}$ $\boldsymbol{h}: \mathbb{R}^n \to \mathbb{R}^m$ $f, \boldsymbol{h} \in$ C^2。

  $$l(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^T \boldsymbol{h}(\boldsymbol{x}) = f(\boldsymbol{x}) + \lambda_1 h_1(\boldsymbol{x}) + \cdots + \lambda_m h_m(\boldsymbol{x})$$

- $\boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda})$ $l(\boldsymbol{x}, \boldsymbol{\lambda})$ $\boldsymbol{x}$

  $$\boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \boldsymbol{F}(\boldsymbol{x}) + \lambda_1 \boldsymbol{H}_1(\boldsymbol{x}) + \cdots + \lambda_m \boldsymbol{H}_m(\boldsymbol{x})$$

  $\boldsymbol{F}(\boldsymbol{x})$ f $\boldsymbol{x}$ $\boldsymbol{H}_k(\boldsymbol{x})$ $h_k, k = 1, \cdots, m$ $\boldsymbol{x}$

$$\boldsymbol{H}_k(x) = \begin{bmatrix} \dfrac{\partial^2 h_k}{\partial x_1^2}(\boldsymbol{x}) & \cdots & \dfrac{\partial^2 h_k}{\partial x_n \partial x_1}(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 h_k}{\partial x_1 \partial x_n}(\boldsymbol{x}) & \cdots & \dfrac{\partial^2 h_k}{\partial^2 x_n^2}(\boldsymbol{x}) \end{bmatrix}$$

5.5

❑ $\boldsymbol{L}(\boldsymbol{x},\boldsymbol{\lambda})$ $l(\boldsymbol{x},\boldsymbol{\lambda}) = f(\boldsymbol{x}) + \lambda^T h(\boldsymbol{x})$ x

$$[\boldsymbol{\lambda H}(\boldsymbol{x})] = \lambda_1 H_1(\boldsymbol{x}) + \cdots + \lambda_m H_m(\boldsymbol{x})$$

$$\boldsymbol{L}(\boldsymbol{x},\lambda) = \boldsymbol{F}(\boldsymbol{x}) + [\boldsymbol{\lambda H}(\boldsymbol{x})]$$

定理5.13（二阶必要条件） $\boldsymbol{x}^*$ $f: \mathbb{R}^n \to \mathbb{R}$ $\boldsymbol{h}(\boldsymbol{x}) = 0$, $\boldsymbol{h}: \mathbb{R}^n \to \mathbb{R}^m, \mathrm{m} \le n, f, \boldsymbol{h} \in C^2$ $\boldsymbol{x}^*$

$\boldsymbol{\lambda}^* \in \mathbb{R}$

1. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) = 0^T$
2. $\boldsymbol{y} \in T(\boldsymbol{x}^*)$, $\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)\boldsymbol{y} \ge 0$

定理5.14 (二阶充分条件) $f, h \in C^2$ $\boldsymbol{x}^* \epsilon \mathbb{R}^n$ $\boldsymbol{\lambda}^* \epsilon \mathbb{R}^n$

1. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) = 0^T$
2. $\boldsymbol{y} \in T(\boldsymbol{x}^*), \boldsymbol{y} \ne 0$, $\boldsymbol{y}^T L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)\mathbf{y} > 0$

$\boldsymbol{x}^*$ f $\boldsymbol{h}(\boldsymbol{x}) = 0$

5.5

❑

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ & \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} \end{aligned}$$

$$f\colon \mathbb{R}^n \to \mathbb{R} \quad \boldsymbol{h}\colon \mathbb{R}^n \to \mathbb{R}^m \quad m \leq n \quad \boldsymbol{g}\colon \mathbb{R}^n \to \mathbb{R}^p$$

定义5.13 $g_j(\boldsymbol{x}) \leq 0$ $\boldsymbol{x}^*$ $g_j(\boldsymbol{x}^*) =$

0 $\boldsymbol{x}^*$ $\boldsymbol{x}^*$

$g_j(\boldsymbol{x}^*) < 0$ $\boldsymbol{x}^*$

5.5

定理5.15 (KKT条件) $f, \boldsymbol{h}, \boldsymbol{g} \in C^l$ $\boldsymbol{x}^*$ $\boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}$ $\boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}$ $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ $\boldsymbol{\mu}^* \in \mathbb{R}^p$

1. $\boldsymbol{\mu}^* \geq \mathbf{0}$
2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\mathrm{T}} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*\mathrm{T}} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}$
3. $\boldsymbol{\mu}^{*\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}^*) = 0$

$\boldsymbol{\lambda}^*$ $\boldsymbol{\mu}^*$ KKT $\boldsymbol{\lambda}^*$ $\boldsymbol{\mu}^*$ KKT

5.5

❑

$$\boldsymbol{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{F}(\boldsymbol{x}) + [\boldsymbol{\lambda}\boldsymbol{H}(\boldsymbol{x})] + [\boldsymbol{\mu}\boldsymbol{G}(\boldsymbol{x})]$$

$\boldsymbol{F}(\boldsymbol{x})$ f $\boldsymbol{x}$ $[\boldsymbol{\lambda}\boldsymbol{H}(\boldsymbol{x})]$

$[\boldsymbol{\lambda}\boldsymbol{H}(\boldsymbol{x})] = \lambda_1\boldsymbol{H}_1(\boldsymbol{x}) + \cdots + \lambda_m\boldsymbol{H}_m(\boldsymbol{x})$ $[\boldsymbol{\mu}\boldsymbol{G}(\boldsymbol{x})]$ $[\boldsymbol{\mu}\boldsymbol{G}(\boldsymbol{x})] = \mu_1\boldsymbol{G}_1(\boldsymbol{x}) + \cdots + \mu_p\boldsymbol{G}_p(\boldsymbol{x})$

$$T(\boldsymbol{x}^*) = \{\boldsymbol{y} \in \mathbb{R}^n : D\boldsymbol{h}(\boldsymbol{x}^*)\boldsymbol{y} = \boldsymbol{0}, Dg_j(\boldsymbol{x}^*)\boldsymbol{y} = 0, j \in J(\boldsymbol{x}^*)\}$$

定理5.16 (二阶必要条件) $\boldsymbol{x}^*$ $f: \mathbb{R}^n \to \mathbb{R}^m$ $\boldsymbol{h}(x) = \boldsymbol{0}$ $\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}$ $\boldsymbol{h}: \mathbb{R}^n \to \mathbb{R}^m (m \leq n)$ $\boldsymbol{g}: \mathbb{R}^n \to \mathbb{R}^p$ $f, \boldsymbol{h}, \boldsymbol{g} \in C^2$ $\boldsymbol{x}^*$ $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ $\boldsymbol{\mu}^* \in \mathbb{R}^P$

1. $\boldsymbol{\mu}^* \geqslant \boldsymbol{0}, Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{\cdot T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*\top} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}, \boldsymbol{\mu}^{*\top}\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^\top$
2. $\boldsymbol{y} \in T(\boldsymbol{x}^*)$ $\boldsymbol{y}^\top \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)\boldsymbol{y} \geqslant 0$

5.5

❑

$$\tilde{T}(\boldsymbol{x}^*,\boldsymbol{\mu}^*) = \{\boldsymbol{y}: D\boldsymbol{h}(x^*)^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}, Dg_i(\boldsymbol{x}^*)\boldsymbol{y} = \boldsymbol{0}^{\top}, i \in \tilde{J}(\boldsymbol{x}^*,\boldsymbol{\mu}^*)\}$$

$\tilde{J}(\boldsymbol{x}^*,\boldsymbol{\mu}^*) = \{i: g_i(\boldsymbol{x}^*) = 0, u_i^* > 0\}$ $\tilde{J}(\boldsymbol{x}^*,\boldsymbol{\mu}^*)$ $\tilde{J}(\boldsymbol{x}^*)$

$\tilde{J}(\boldsymbol{x}^*,\boldsymbol{\mu}^*) \subset \tilde{J}(\boldsymbol{x}^*)$ $T(\boldsymbol{x}^*)$ $\tilde{T}(\boldsymbol{x}^*,\boldsymbol{\mu}^*)$

$T(\boldsymbol{x}^*) \subset \tilde{T}(\boldsymbol{x}^*,\boldsymbol{\mu}^*)$

定理6.3 (二阶充分条件) $f,\boldsymbol{g},\boldsymbol{h} \in C^2$ $\boldsymbol{x}^* \in \mathbb{R}^m$

$\boldsymbol{\lambda}^* \in \mathbb{R}^m$ $\boldsymbol{\mu}^* \in \mathbb{R}^p$

1. $\boldsymbol{\mu}^* \geqslant \boldsymbol{0}, Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\top} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*\top} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^{\top}, \boldsymbol{\mu}^{*\top} g(\boldsymbol{x}^*) = 0$
2. $\boldsymbol{y} \in \tilde{T}(\boldsymbol{x}^*,\boldsymbol{\mu}^*)$ $\boldsymbol{y} \neq 0$ $\boldsymbol{y}^{\top}\boldsymbol{L}(\boldsymbol{x}^*,\boldsymbol{\lambda}^*;\boldsymbol{\mu}^*)\boldsymbol{y} > 0$

$\boldsymbol{x}^*$ $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}$